OPTIMAL DECISION MAKING
MGT-483

GROUP Q

GROUP PROJECT: OPTIMAL TRANSPORT AND DISTRIBUTIONALLY ROBUST OPTIMIZATION

May 12, 2019

Feng, Wentao
Sun, Zhaodong
Wang, Yunbei
Wu, Yi-Shiun

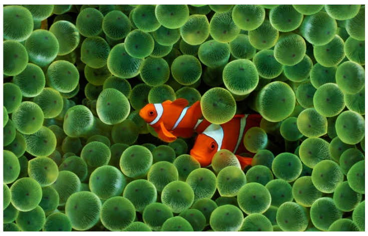
Contents

1 Computing Wasserstein Distances

The Wasserstein distance between $\mathbb{P}$ and $\mathbb{Q}$ is define as the square root of the optimal value. In this case the optimal value we obtained is 0.2667 and the Wasserstein distance is 0.5164.

2 Color Transfer

In this section, we performed color transformation between two given RGB images by solving a transportation problem, which seeks the most natural way of transporting the color distribution of one image to that of another one. The color transfer algorithm imposes a palette of the target image on the source image and preserves the geometry of the source image.



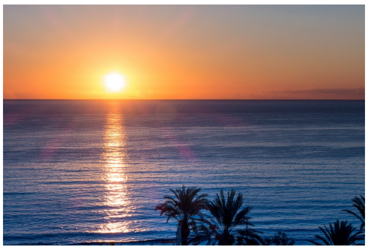
(a) fish

(b) coral

(c) spring

Figure 1: Original source images.

(a) view



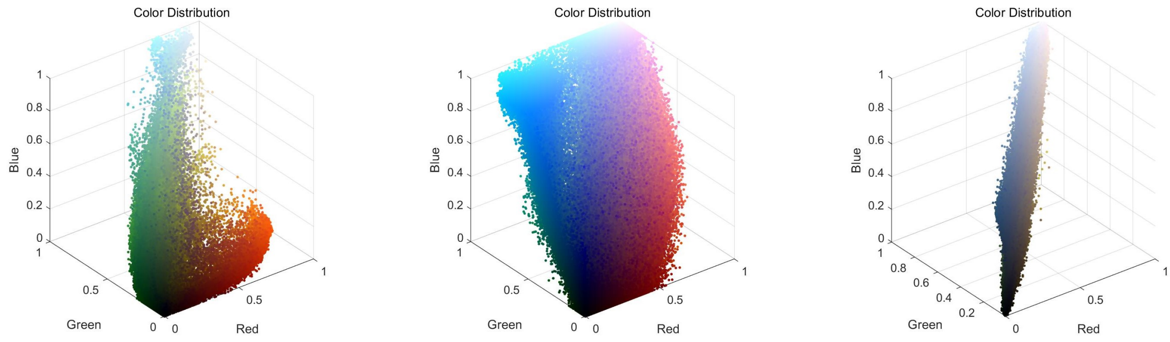
(b) sunset

(c) fall

Figure 2: Target images of re-coloring.

2.1 Color distributions

Each pixel in the picture is considered as a point in the 3-dimensional RGB color space. We reshape the original image into a 2D array which each column is the RGB values of a single pixel. With the *scatter3* function in MATLAB, we are able to plot the color distribution of source images(Figure 3).



(a) fish **(b)** coral **(c)** spring

Figure 3: Color distribution of three pictures

2.2 Formulation

$$\min_{\pi\in\mathbb{R}_+^{(W1\cdot H1)\times(W2\cdot H2)}} \sum_{i=1}^{W_1\times H_1}\sum_{j=1}^{W_2\times H_2} c_{ij}\pi_{ij} \tag{1}$$

Subject to

$$\sum_{i=1}^{W_1\times H_1}\pi_{ij}=\frac{1}{W_2\times H_2}, \forall j=1,\dots,W_2\times H_2 \tag{2}$$

$$\sum_{j=1}^{W_2\times H_2}\pi_{ij}=\frac{1}{W_1\times H_1}, \forall i=1,\dots,W_1\times H_1 \tag{3}$$

$$\pi_{ij}\geq 0, \forall i=1,\dots,W_1\times H_1, \forall j=1,\dots,W_2\times H_2 \tag{4}$$

where

$$c_{ij}=\| x_i-y_j \|_2^2$$

$$X^{'}=\begin{bmatrix} x_1^T \\ \vdots \\ x_i^T \\ \vdots \\ x_{W1\times H1}^T \end{bmatrix} \quad Y^{'}=\begin{bmatrix} y_1^T \\ \vdots \\ y_j^T \\ \vdots \\ y_{W2\times H2}^T \end{bmatrix}$$

The number of variables is equal to $W_1\times H_1\times W_2\times H_2$. The number of constraints is equal to $W_2\times H_2+W_1\times H_1+W_1\times H_1\times W_2\times H_2$.

2.3 Re-coloring images

With the color map we obtain by solving the optimization problem above with sampled source image. Figure 4 shows the full re-colored source images based on the color palette of their designated target images (Figure2).

(a) fish **(b)** coral **(c)** spring

Figure 4: Re-colored source images based on target images.

2.4 Bonus

The $\pi^\star$ is the optimal transportation map for source image to re-color based on the color palette of target images. The transpose matrix ${\pi^\star}^T$ is the reverse transportation map and it can re-color target image based on the color palette of source images. Using the same method in section 2.3, we are able to re-color the target images by their correspond source images.

(a) view **(b)** sunset **(c)** fall

Figure 5: Re-colored target images based on source images

3 Data-Driven Portfolio Optimization

3.1 Formulation of SAA

$$\begin{aligned} &\max_{x \in \chi} \frac{1}{N} \sum_{i=1}^{N} u(x^\top \xi_i) && (5) \\ \text{subject to:} \quad & u(x^\top \xi_i) = \min_{l=1,\dots,L} \{a_l x^\top \xi_i + b_l\} \\ & \mathbb{1}^\top x = 1 \\ & x \geq 0 \end{aligned}$$

We can reformulate the constraint and the get the following representation:

$$\begin{aligned} &\max_{x \in \chi} \frac{1}{N} \sum_{i=1}^{N} t_i && (6) \\ \text{subject to:} \quad & t_i \leq a_l x^\top \xi_i + b_l, \quad \forall l = 1, \dots, l, \quad \forall i = 1, \dots, N \\ & \mathbb{1}^\top x = 1 \\ & x \geq 0 \end{aligned}$$

3.2 Formulation of DRO

The minimization part of out objective function can be re-write in the following form:

$$\begin{aligned} &\min_{\mathbb{Q}_i \geq 0} \sum_{i=1}^{N} \mathbb{Q}_i u(x^\top \xi_i) && (7) \\ \text{subject to:} \quad & d(\hat{\mathbb{P}}, \mathbb{Q}) \leq \rho \end{aligned}$$

Specifically, $d(\hat{\mathbb{P}}, \mathbb{Q})$ is the square root of the optimal value of the following optimization problem:

$$
\begin{aligned}
\min_{\pi_{ji}} \sum_{j=1}^{N} \sum_{i=1}^{N} ||\xi_j - \xi_i||^2 \pi_{ji} \qquad (8)\\
\text{subject to:} \quad \sum_{i=1}^{N} \pi_{ji} = \frac{1}{N}, \quad \forall j\\
\sum_{j=1}^{N} \pi_{ji} = \mathbb{Q}_i, \quad \forall i\\
\pi_{ji} \geq 0
\end{aligned}
$$

Now, we eliminate the min operator with robust optimization trick in convex optimization. So, we transform our problem into following formation:

$$
\begin{aligned}
\min_{\mathbb{Q}_i \geq 0, \pi_{ji}} \sum_{i=1}^{N} \mathbb{Q}_i u(x^\top \xi_i) \qquad (9)\\
\text{subject to:} \quad \sum_{j=1}^{N} \sum_{i=1}^{N} ||\xi_j - \xi_i||^2 \pi_{ji} \leq \rho^2\\
\sum_{i=1}^{N} \pi_{ji} = \frac{1}{N}, \quad \forall j\\
\sum_{j=1}^{N} \pi_{ji} = \mathbb{Q}_i, \quad \forall i\\
\pi_{ji} \geq 0
\end{aligned}
$$

We can replace $\mathbb{Q}_i$ with the sum of π_{ji}:

$$
\begin{aligned}
\min_{\pi_{ji}} \sum_{i=1}^{N} \sum_{j=1}^{N} \pi_{ji} u(x^\top \xi_i) \qquad (10)\\
\text{subject to:} \quad \sum_{j=1}^{N} \sum_{i=1}^{N} ||\xi_j - \xi_i||^2 \pi_{ji} \leq \rho^2\\
\sum_{i=1}^{N} \pi_{ji} = \frac{1}{N}, \quad \forall j\\
\pi_{ji} \geq 0 \quad \forall i, j
\end{aligned}
$$

Through the strong duality of linear programming, we get the duality problem:

$$\begin{aligned} &\max_{\lambda, s_j} \lambda\rho^2 + \frac{1}{N}\sum_{j=1}^{N} s_j \\ \text{subject to:} \quad & ||\xi_j - \xi_i||^2 \lambda + s_j \leq u(x^\top \xi_i) \quad \forall i, j \\ & \lambda \leq 0 \end{aligned} \tag{11}$$

Then, we add the outer part of our original objective function:

$$\begin{aligned} &\max_{\lambda, s_j, x} \lambda\rho^2 + \frac{1}{N}\sum_{j=1}^{N} s_j \\ \text{subject to:} \quad & ||\xi_j - \xi_i||^2 \lambda + s_j \leq a_l x^\top \xi_i + b_l \quad \forall i, j, l \\ & \lambda \leq 0, \quad \mathbb{1}^\top x = 1, \quad x \geq 0 \end{aligned} \tag{12}$$

3.3 SAA implementation

With $N = 10,000$, the optimal value is 0.5712.

3.4 DRO implementation

With 30 training sample, 10,000 test sample and a Wasserstein radius of $\rho = 0.9$, the mean out-of-sample utility is 0.5255

3.5 Independent datasets experiment

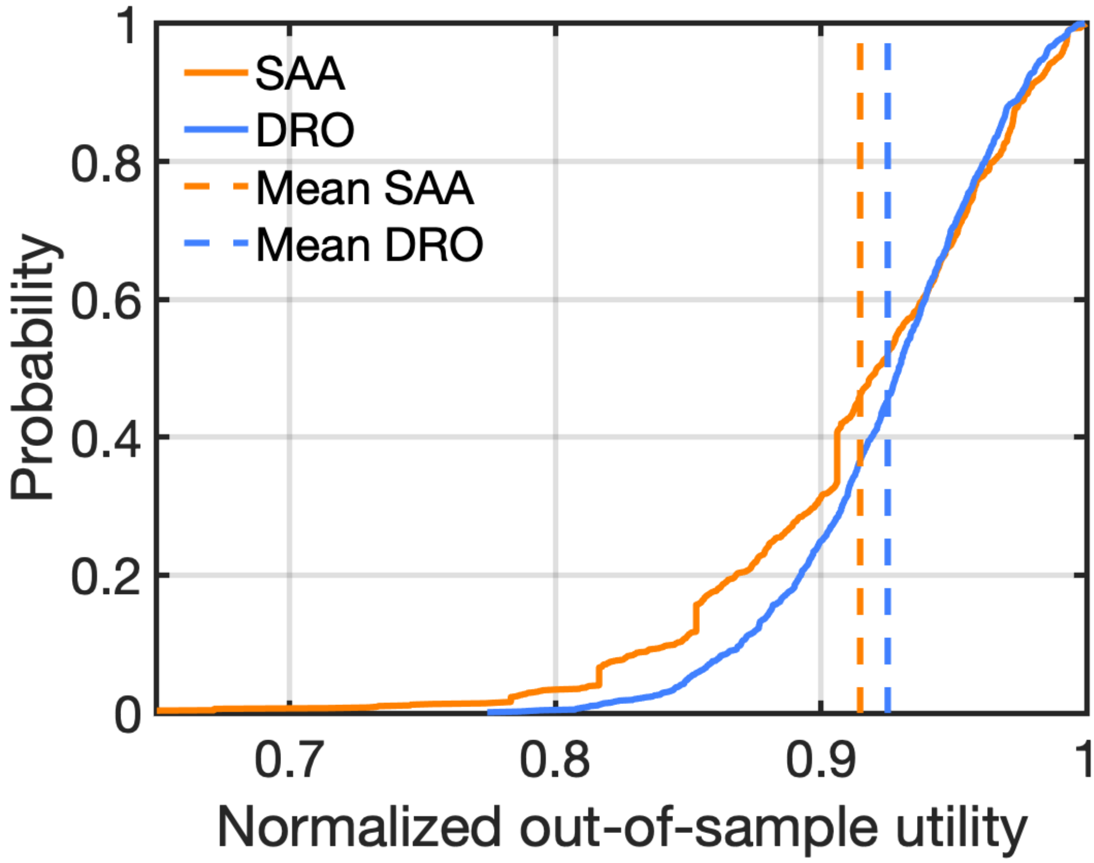


Figure 6: Cumulative distribution of 1,000 independent datasets experiment

3.6 Interpretation

Note: all values we discuss are normalized concerning the result of Exercise 3.3.

- The DRO has a better average performance than the SAA. In our 1000 independent datasets, the average normalized out-of-sample utility is 0.91 +/- 0.06 for the SAA and 0.93 +/- 0.04 for the DRO. In other words, the results of the DRO have better average performance and less variation at the same time.
- The DRO is more reliable. With the probability of 95%, the DRO can get performance better than 0.85, and the SAA can only get 0.82. That is also why the orange curve is above the blue curve in the major part of the graph from the beginning.

- The SAA is more likely to achieve excellent performance. With the probability of 5%, the SAA can get performance better than 0.99, and the DRO can only get 0.98. That is the reason the blue curve is above the orange one during the tail part.